

AD_____

Grant Number DAMD17-94-J-4508.

TITLE: California Cancer Registry Enhancement for Breast Cancer Research

PRINCIPAL INVESTIGATOR: William Wright, Ph.D.

CONTRACTING ORGANIZATION: California Public Health Foundation
Berkeley, California 94704-1103

REPORT DATE: October 1997

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19980618 125

DTIC QUALITY INSPECTED 1

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE October 1997	3. REPORT TYPE AND DATES COVERED Annual (1 Oct 96 - 30 Sep 97)	
4. TITLE AND SUBTITLE California Cancer Registry Enhancement for Breast Cancer Research			5. FUNDING NUMBERS DAMD17-94-J-4508	
6. AUTHOR(S) William Wright, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) California Public Health Foundation Berkeley, California 94704-1103			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200) The purpose of this project is to enhance the value of the California Cancer Registry as a research tool for clinicians and epidemiologists interested in conducting breast cancer research. The objectives of the project are to: (1) classify breast cancers according to the major staging schemes currently in use in the U.S.; (2) increase the amount of treatment data for breast cancer; and (3) link breast cancer case data with other data bases to improve survival information and collect co-morbidity information. Difficulties in finding qualified registrars to perform EOD coding have been overcome and all cases from 1988 forward have been coded. Computer software to process correction records has been developed and activated. Physician follow back for treatment data has been initiated in one region. The scope of this activity was expanded from early stage breast cancers to all cancers. Resources to conduct this activity are much greater than had been anticipated. Linkage activities are on schedule. The project has generated one oral and one poster session presentation at professional meetings, and two papers are in preparation.				
14. SUBJECT TERMS Breast Cancer			15. NUMBER OF PAGES 77	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

✓ _____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

_____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

_____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

William E. Wright 2/9/88
PI - Signature Date

TABLE OF CONTENTS

	<u>Page</u>
Front Cover	1
Report Documentation Page	2
Foreword	3
Introduction	5
Body	5
Conclusions	24
References	25
Appendixes	
I. Correction Record Processing Specifications	26
II. Female Breast Cancer Treatment Groups	47
III. Capture of Missing First Course of Therapy	50
IV. Cover Letter and Physician Questionnaire	56
V. Poster Presentation	59

INTRODUCTION

The purpose of this project is to enhance the value of the California Cancer Registry (CCR) as a research tool for clinicians and epidemiologists interested in conducting breast cancer research. The goals are to code in greater detail the extent of disease at the time of diagnosis, to gather complete information about the first course of treatment, to collect follow-up information about vital status, to code information about occupation and industry, to link the CCR files with a variety of existing files containing information on patterns and costs of care, and to develop mechanisms by which a broad audience of breast cancer researchers can obtain access to the CCR database.

BODY

Progress to date:

Objective 1 - Code SEER Extent of Disease for all breast cancers diagnosed in California starting with January 1, 1998.

Between 1988 and 1993 all breast cancers were staged according to the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program Summary Staging Guide (1), basically a classification of cases into *in situ*, localized, regional, and distant disease. In 1994 the CCR changed its reporting requirement from the SEER Summary Stage to the SEER Program's Extent of Disease (EOD) (2) classification scheme in order to be able to apply a computer program available from the NCI to classify breast cancer cases into the TNM classifications and the Staging Categories (0, I, II, III, IV) of the American Joint Committee on Cancer (3). A major objective of this award has been to reclassify all breast cancer cases diagnosed between 1988-1993 according to the SEER EOD classification scheme.

Objective 1 was completed this past year. For the purposes of cancer reporting, California has been divided into 10 geographical regions. Table 1 shows the number of female breast cancer cases by region and year of diagnosis (1988-1995) which have been received by the CCR as of October 1997. Extent of Disease is categorized by three fields: Extension of the tumor (DIREXTTU), tumor size, (TSIZETU), and lymph node involvement (LNSUMTU). A code of "unknown" means that the patient medical record did not contain sufficient information to code this field while a code of "blank" means that the record was not searched for this variable. Tables 2-4 present the number of cases which are coded as "unknown" or "blank" for each EOD field by region and year of diagnosis, and Figures 1-10 present the percent of cases which are coded as "unknown" or "blank" for each EOD field by reporting region and year of diagnosis. Statewide, among 160,809 records only 59 records are coded "blank" on the DIREXTTU variable, 26 are "blank" on TSIZETU, and 54 are "blank" on LNSUMTU. DIREXTTU is recorded as missing on 6,135 (3.8%) cases, TSIZETU is missing on 24,841 (15.4%), and LNSUMTU is missing on 14,746 (9.2%).

The CCR has obtained the computer software from the NCI for classifying breast cancer cases into the TNM staging categories. The final report will contain an analysis of trends in breast cancer stage at diagnosis by year.

Table 1

Female breast cancer cases (in situ and invasive), resident within
region, CCR (Jan98), 1988-1995

TABLE OF REGIONID BY YEARDX

REGIONID(Reporting Region)	YEARDX(Year of diagnosis (YY))								Total
Frequency	88	89	90	91	92	93	94	95	
1	1278	1330	1291	1412	1438	1443	1432	1601	11225
2	1175	1120	1321	1367	1372	1397	1491	1597	10840
3	1737	1689	1812	1888	1967	1842	1901	1986	14822
4	829	784	852	896	911	926	960	943	7101
5	1473	1434	1650	1680	1676	1785	1672	1770	13140
6	974	962	1012	1087	1147	1087	1216	1102	8587
7	1535	1583	1753	1871	1962	1879	1859	1890	14332
8	2883	2848	2981	3019	3086	2986	3157	3273	24233
9	5483	5172	5280	5244	5510	5341	5386	5487	42903
10	1611	1530	1661	1651	1821	1788	1736	1828	13626
Total	18978	18452	19613	20115	20890	20474	20810	21477	160809

Table 2

EOD coding on resident female breast cancer (in situ and invasive) cases on Jan98 submission
 Number and percent of cases where DIREXTTU is coded as unknown (9s) or blank

Cases		Year of Dx							
		88	89	90	91	92	93	94	95
Region	Coding								
1	Unknown	63	69	69	48	47	69	34	32
2	Blank	3	3	3	0	0	0	0	0
	Unknown	79	52	48	46	60	75	66	56
3	Blank	0	0	2	1	0	0	0	0
	Unknown	54	50	47	60	47	51	56	40
4	Blank	0	1	0	1	1	0	0	0
	Unknown	40	36	37	32	42	25	36	24
5	Blank	1	1	1	1	2	1	0	0
	Unknown	45	42	62	54	75	69	53	77
6	Blank	0	0	0	0	0	1	0	0
	Unknown	52	52	58	33	30	54	80	45
7	Blank	3	5	6	0	0	3	0	0
	Unknown	87	89	94	93	112	100	77	90
8	Unknown	81	80	92	81	92	101	56	62
9	Blank	2	4	4	4	0	0	0	0
	Unknown	261	238	224	191	174	181	182	152
10	Blank	1	1	3	0	0	0	0	0
	Unknown	93	94	104	91	95	84	44	69

Table 3

EOD coding on resident female breast cancer (in situ and invasive) cases on Jan98 submission
 Number and percent of cases where TSIZETTU is coded as unknown (9s) or blank

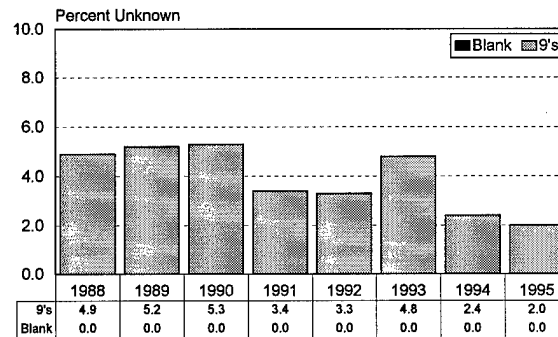
Cases		Year of Dx							
		88	89	90	91	92	93	94	95
Region	Coding								
1	Unknown	283	255	255	265	274	282	235	237
2	Blank	0	0	2	0	0	0	0	0
	Unknown	235	241	271	257	260	204	245	202
3	Unknown	423	372	398	343	383	216	240	234
4	Unknown	132	144	165	168	176	158	155	161
5	Blank	0	0	0	0	0	1	0	0
	Unknown	284	242	282	295	316	314	228	253
6	Unknown	266	258	258	216	204	221	240	165
7	Blank	3	2	5	0	0	2	0	0
	Unknown	226	231	330	299	290	262	266	264
8	Unknown	504	420	473	420	415	383	377	396
9	Blank	2	1	2	1	0	0	0	0
	Unknown	813	716	641	692	660	679	631	535
10	Blank	1	1	3	0	0	0	0	0
	Unknown	265	273	309	283	228	218	221	210

Table 4

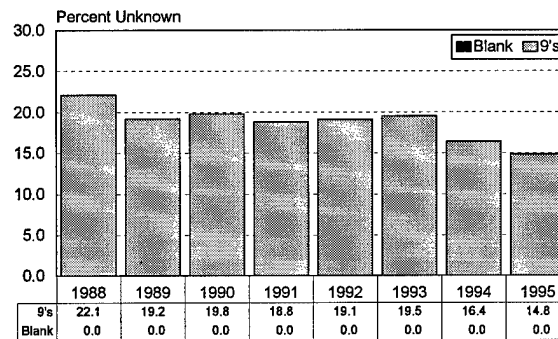
EOD coding on resident female breast cancer (in situ and invasive) cases on Jan98 submission
 Number and percent of cases where LNSUMTU is coded as unknown (9s) or blank

Cases		Year of Dx							
		88	89	90	91	92	93	94	95
Region	Coding								
1	Unknown	108	149	123	116	116	181	144	161
2	Blank	2	4	3	0	1	0	0	0
	Unknown	141	135	162	148	184	190	136	131
3	Blank	0	0	0	1	0	0	0	0
	Unknown	206	188	233	142	275	165	130	128
4	Blank	0	1	0	1	0	1	0	0
	Unknown	85	92	81	98	111	101	87	74
5	Blank	2	1	1	1	2	1	0	0
	Unknown	121	123	174	143	163	180	187	177
6	Unknown	122	136	150	127	129	177	157	117
7	Blank	3	4	5	0	0	2	0	0
	Unknown	95	130	134	158	162	171	148	176
8	Unknown	278	250	227	208	228	275	270	291
9	Blank	2	3	4	4	0	0	0	0
	Unknown	545	415	406	405	472	436	444	418
10	Blank	1	1	3	0	0	0	0	0
	Unknown	140	162	165	136	142	143	98	114

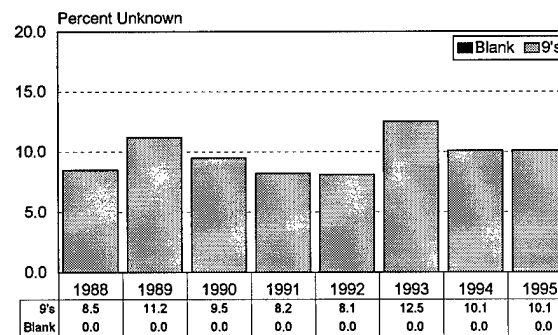
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 1



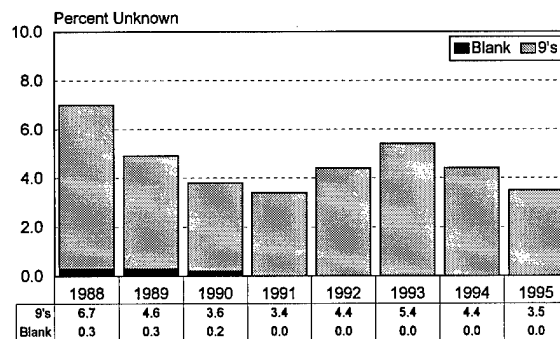
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 1



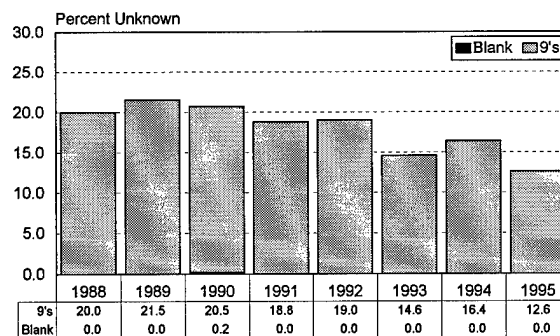
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 1



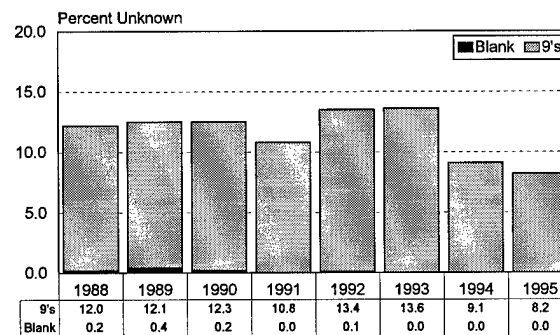
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 2



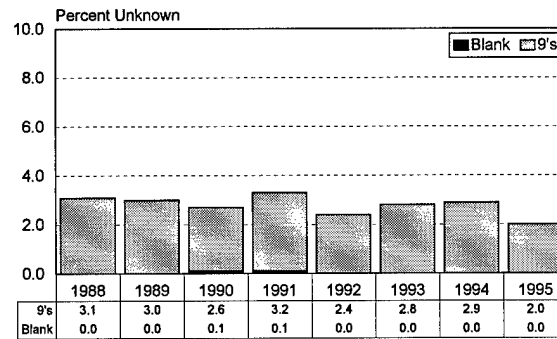
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 2



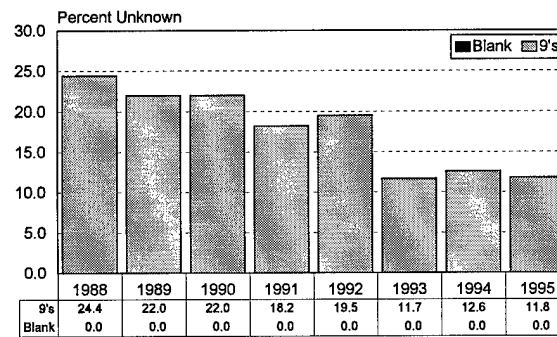
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 2



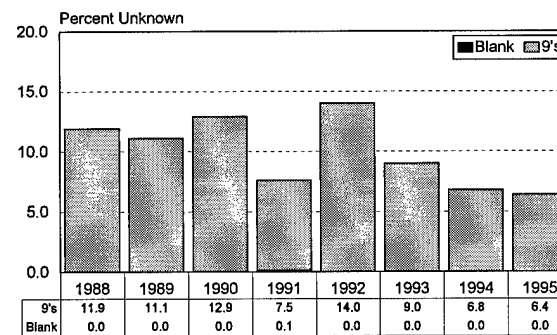
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 3



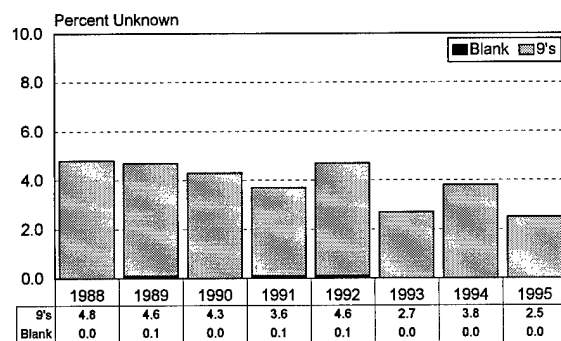
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZEITU Coded 9's or Blank
Region 3



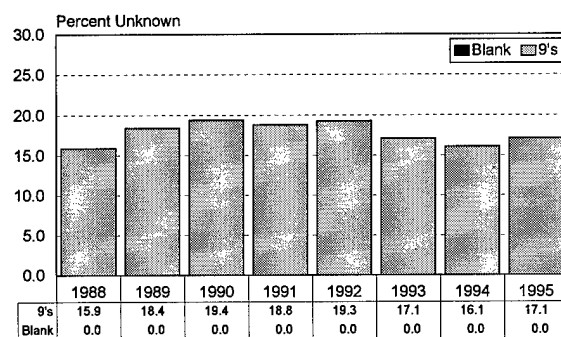
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 3



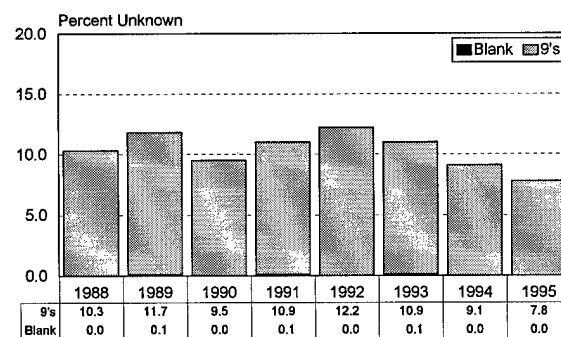
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 4



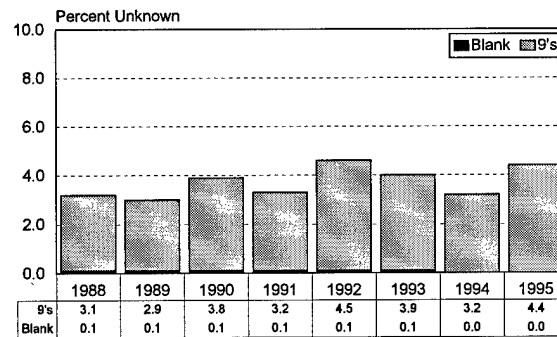
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZEITU Coded 9's or Blank
Region 4



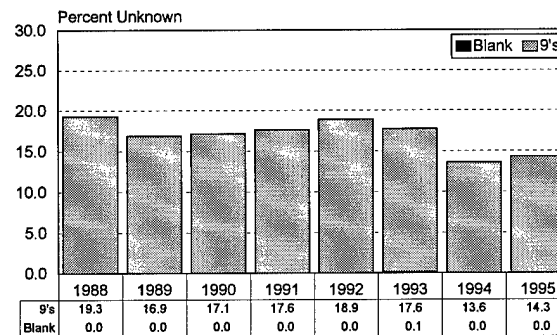
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 4



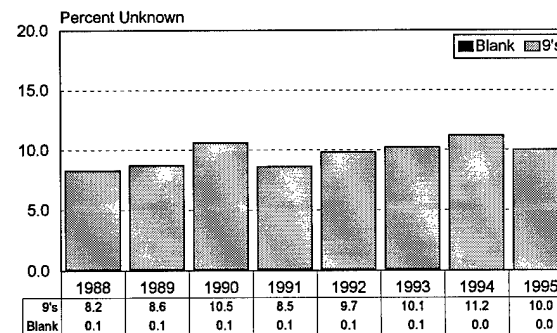
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 5



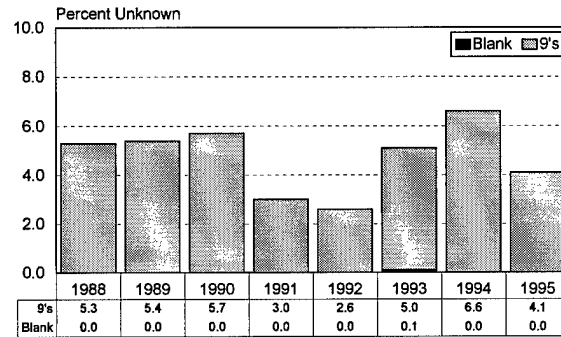
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZEU Coded 9's or Blank
Region 5



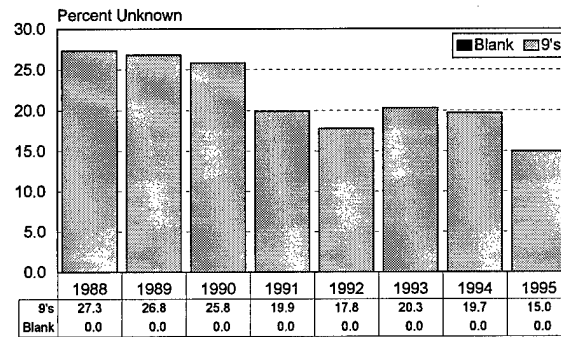
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 5



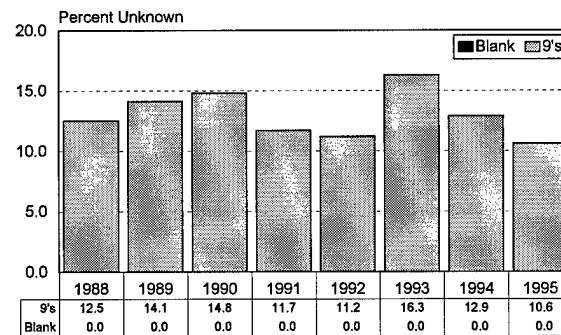
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 6



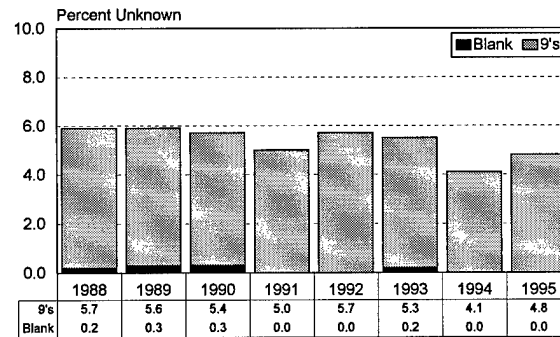
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 6



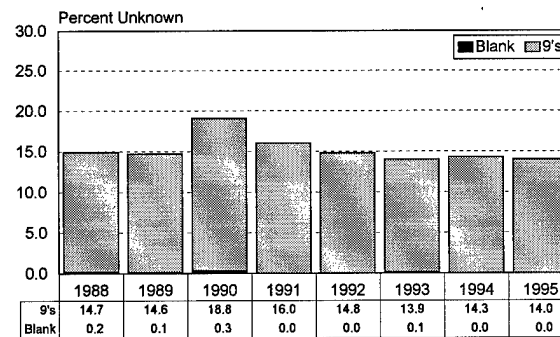
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 6



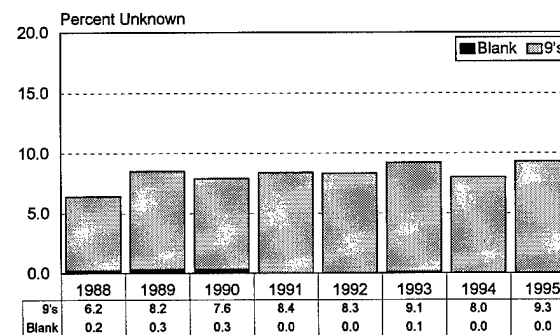
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 7



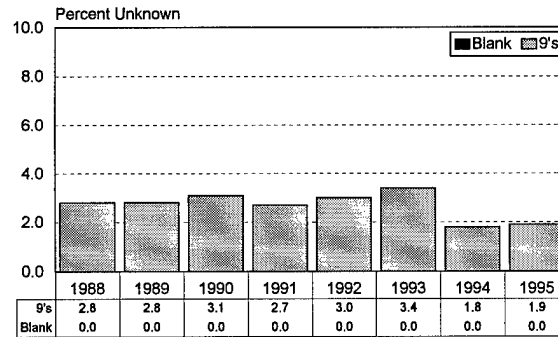
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 7



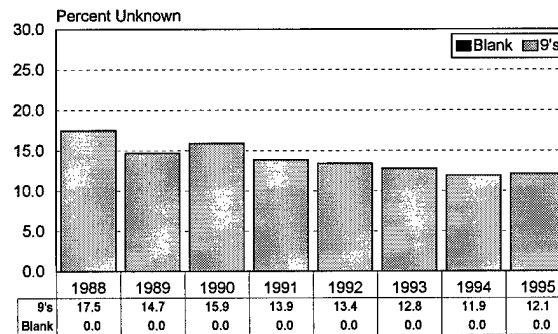
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 7



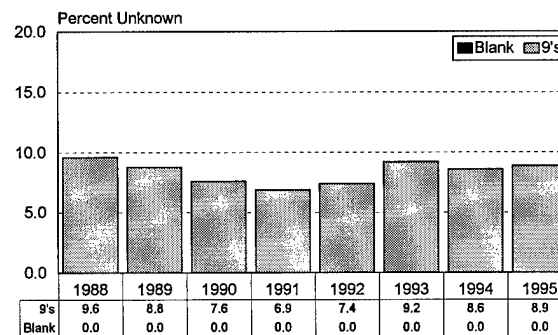
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 8



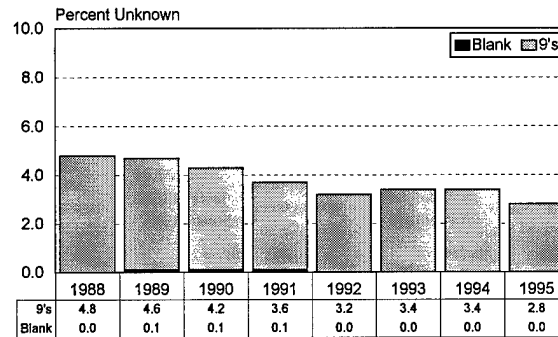
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 8



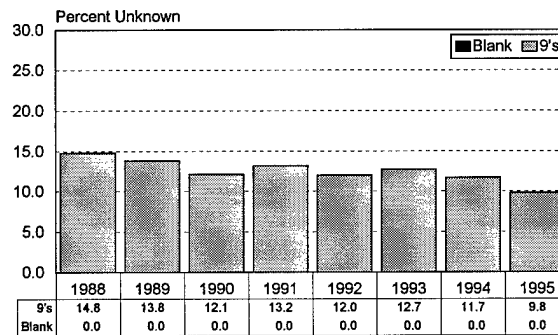
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 8



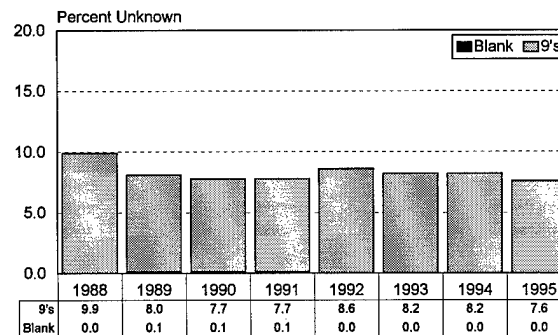
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 9



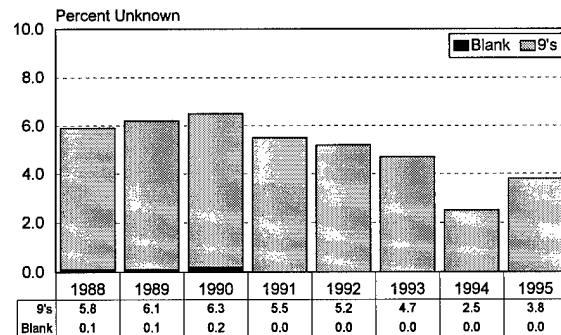
Breast Cancer EOD Coding Status, Jan 98
Percent TSIZETTU Coded 9's or Blank
Region 9



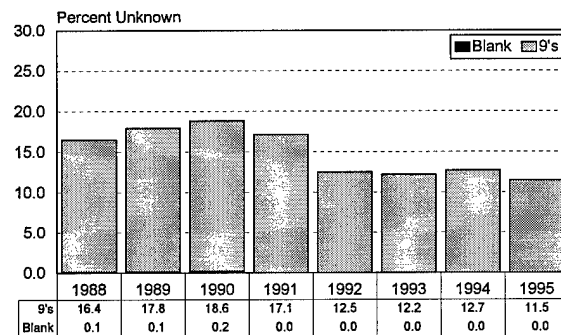
Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 9



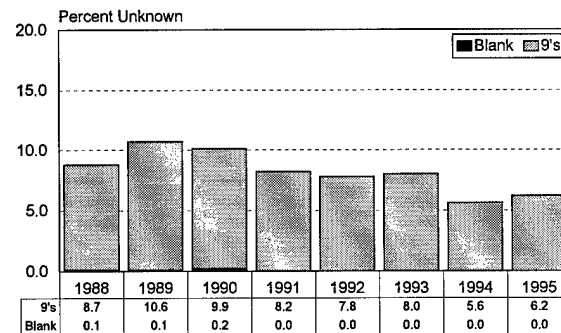
Breast Cancer EOD Coding Status, Jan 98
Percent DIREXTTU Coded 9's or Blank
Region 10



Breast Cancer EOD Coding Status, Jan 98
Percent TSIZEU Coded 9's or Blank
Region 10



Breast Cancer EOD Coding Status, Jan 98
Percent LNSUMTU Coded 9's or Blank
Region 10



Objective 2 - Collect complete first course of treatment information for all breast cancers diagnosed from 1993 through 1997.

Until recently, most population-based registries outside the SEER Program have been incidence only registries and have not been concerned with the collection of treatment data. Since its inception, the CCR has recorded the first course of cancer treatment for all patients. Unfortunately, the data are known to be incomplete, especially for those cancer sites such as breast cancer which may be treated with a first course of chemotherapy and eventually followed up with radiation therapy. Chemotherapy and radiation therapy are primarily given outside the acute care hospital setting, and hospital medical records often lack the details of the complete first course of therapy that was given.

Data on female breast cancers come in to the CCR from multiple sources. Frequently there are admissions to more than one hospital, and additional treatment information may come in from a physicians office or from an updated hospital record. When a new patient record is received from a hospital by a regional registry, it is either entered as a new case or "consolidated" with the records from other facilities into a single record. Hospitals frequently abstract and report a case shortly after it is diagnosed and treated in that facility. Subsequently, the hospital registrar may learn of additional treatment and update the hospital record. The updated information is then transmitted to the regional registry as a "correction" record.

As stated in the prior progress report, due to limited resources the CCR had not developed software to process these correction records before initiating this project. Only Region 8, the San Francisco Bay Area Registry, had processed these records on a routine basis using resources available from the SEER Program, and they had to process them manually. The other regional registries in California had stockpiled their correction records since the implementation of statewide reporting in 1988. Consequently there was an unknown amount of treatment information contained in the stockpiled correction records. This information needed to be processed and added to the main data base before any given breast cancer record could be compared with the standard recommended treatment, and before any routine follow-back to physicians concerning possible incomplete treatment could be initiated.

Last year the CCR's Correction Records Processing Task Force completed the work of developing specifications for comparing correction record data with the main data base, developing decision rules for handling discrepancies and for automating as much of the process as possible. Appendix I contains a copy of the final processing specifications. Computer software for processing correction records has been written, tested, and installed in the four different software systems used by the ten regional registries. The backlog of breast cancer correction records is expected to be processed early in the next year of this project. (All of the specification and software development was funded with breast cancer tobacco tax funds that were available to the CCR.)

During the past year Region 8 developed and implemented methods for comparing treatment information contained in the registry file with a treatment standard and, if different, conducting follow back to query physician's offices. The Breast Cancer Treatment Follow Back Protocol consists of: (1) a standard for comparing recommended breast cancer treatment with treatment recorded in the registry record; (2) computer programs to perform the comparison; (3) criteria for excluding cases

from follow back; (4) updating physician addresses; (5) computer programs to generate customized letters to physicians requesting treatment on specific treatment that was recommended but not recorded in the registry record; (6) interaction with hospital cancer registry staff in order to determine who should perform the follow back, i.e. central registry or hospital registry staff; (7) criteria for intensity of follow back, i.e. multiple query letters and/or actual visits by program staff to physician offices to extract information from medical records; and (8) data entry onto "correction" records for processing to update registry data files.

As stated in the Year 01 Progress Report, standard/recommended/state-of-the-art treatment for each stage and type of breast cancer is included in the NCI's Physician Data Query (PDQ) system which is available to all practicing physicians via the Internet (http://cancernet.nci.nih.gov/clinpdq/soa/Breast_cancer_Physician.html) or the NCI's Cancer Information Service (1-800-4-CANCER). Naturally, not all physicians utilize the PDQ, and some physicians do not feel that it is appropriate for NCI to "dictate" how patients should be treated, believing that the choice belongs to the physician and patient. Nevertheless, the comparison standard chosen for this project was the PDQ.

A computer program had been developed earlier by the Seattle SEER Program located at the Fred Hutchinson Cancer Research Center to compare recommended cancer treatment data versus that recorded in a central cancer registry. This program was made available to the CCR as a model. CCR staff reviewed and made minor modifications of the Seattle PDQ classifications. Appendix II contains criteria for the 21 breast cancer treatment groups that result from the PDQ. Staff of the Region 8 Registry developed computer programs to assign cases to the 21 PDQ groups, compare against criteria for excluding cases from follow back, update physician addresses with more current information, and generate customized letters to physicians.

All of the developmental work listed above was not completed until early 1997. Due to the lag time between diagnosis date and treatment follow back (over four years for early 1993 diagnoses) and subsequent difficulty obtaining medical records, we decided to focus the first efforts in Region 8 on 1994 cases. For that year 4,795 female breast cancers had been reported. After the initial review, 905 cases were excluded from follow back. Table 5 presents the number of cases excluded by reason. Of the remaining cases, 1,053 had been diagnosed and treated in the Kaiser Permanente Health Maintenance Organization. Region 8 arranged for Kaiser Permanente to conduct a separate search for missing information utilizing their own extensive computer files. The results of that search are documented in Appendix III. In summary, 336 (7%) of 1994 cases had complete treatment when compared with the PDQ and were excluded from follow back, 569 (12%) were excluded for other reasons, 1,053 (22%) were sent to Kaiser Permanente for follow back via their computer records, and 2,837 (59%) were designated for physician follow back.

A cover letter from the Director of the Northern California Cancer Center (Region 8) and a customized questionnaire (Appendix IV) asking only for treatment that the PDQ recommended but that was missing from the Region 8 case report was prepared and mailed for each of the cases designated for follow back. After considerable effort including second and third mailings, mailings to different physicians than the one listed in the cancer registry report at the "follow up physician", telephone calls, and staff visits to physician offices to abstract the requested information from patient records, 2,302 responses were received for an 81% response rate (Table 5). Among these responses,

Tables

TOTAL CASES IN 1994 PDQ TREATMENT FOLLOW-BACK - BREAST CASES

4795

EXCLUSIONARY CODES

TOTAL FOR STATUS CODE:	01 COMPLETE TX - NO FOLLOW-BACK NEEDED	336
TOTAL FOR STATUS CODE:	02 NON-RES DX	242
TOTAL FOR STATUS CODE:	03 DC ONLY	12
TOTAL FOR STATUS CODE	04 FIRST DX AT AUTOPSY	3
TOTAL FOR STATUS CODE:	05 PHYSICIAN ONLY (HOSP 00803)	35
TOTAL FOR STATUS CODE:	06 CORONER (HOSP 00802)	4
TOTAL FOR STATUS CODE:	07 PATIENT REFUSED ALL TX	1
TOTAL FOR STATUS CODE:	09 HOSPITAL CLOSED	33
TOTAL FOR STATUS CODE:	10 NO CONTACT DOC AVAILABLE	186
TOTAL FOR STATUS CODE:	11 PATIENT EXPIRED	19
TOTAL FOR STATUS CODE:	14 DOCTOR OUT OF REGION	<u>34</u>

TOTAL EXCLUSIONARY CASES

905

TOTAL CASES MAILED TO PHYSICIAN

2837

TOTAL RESPONSES

2302 (81%)

ADDITIONAL TREATMENT OBTAINED	578 (25%)
NO ADDITIONAL TREATMENT OBTAINED	<u>1724</u> (75%)
TOTAL	2302

TOTAL WITH NO RESPONSE FROM PHYSICIAN

535

TOTAL WAITING FOR RESPONSE FROM KAISERS

1053

TOTAL

4795

08/04/97

25% of the cases contained additional treatment information. For the 1,053 Kaiser Permanente cases (Appendix III), 1,039 (98.7%) were linked to Kaiser historical files and additional treatment information was obtained for 153 (14.7%) of the linked records.

In summary, 3,890 Region 8 cases were designated for follow back and additional treatment information was obtained for 731 (18.8%) of those cases. This additional treatment information has been entered onto the main data file via the "correction" record process. Comparison of these 731 cases with the PDQ will be conducted in the next year and the results will be contained in the Final Report.

During the course of implementing the Breast Cancer Treatment Follow Back Protocol for the CCR, we discovered the process was more time sensitive and resource intensive than we had imagined when we initially proposed this effort. Even with additional funds from the NCI and the California Breast Cancer Tobacco Tax Research Fund that are available to the CCR, we cannot perform follow back on all cases from 1993 through 1997. Cases diagnosed in 1993 and 1994 are now too old for their records to be readily available in physician offices. Follow back requires considerably more staff resources than we first estimated due to the necessity for multiple attempts to contact the physician of record, tracing physicians who have moved, interacting with hospitals for coordinating follow back activities that they may be engaged in, and physically going to physician offices to abstract treatment information from their files. Cases diagnosed during calendar year 1997 will not be completely received in CCR Regional Registry offices and processed through our quality control edits and visual review by July 1, 1998 which is our cut-off date for treatment follow back activities in order to complete the follow back and report by the end of the grant period. Consequently, we have modified our Scope of Work to collect first course of treatment information for all breast cancer cases statewide diagnosed only for the time period 1995 through 1996.

Objective 3 - Collect patient follow-up information on all breast cancers diagnosed from 1988 forward by linking the CCR files with Department of Motor Vehicles and voter registration files.

The results of our linkage with the Department of Motor Vehicles (DMV) files were described in our last annual report, and no additional DMV linkages were performed during the past year. The CCR expects to incorporate linkage of its entire data base with the DMV into an annual production process. Linkage with the 1998 DMV file will be conducted in the next year and the results will be presented in the Final Report.

Linkage with voter registration files has not been accomplished. This task would require more resources than are available, and it was deleted during budget negotiations at the beginning of the grant.

Objective 4 - Complete occupation/industry coding for all breast cancer cases from 1998 through 1997.

This objective was deleted during budget negotiations at the beginning of the grant.

Objective 5 - Link CCR files with data from several large breast cancer screening programs to correlate screening status with subsequent diagnostic status.

The CCR is collaborating with the California Breast and Cervical Cancer Control Program (BCCCCP) (funded by the Centers for Disease Control (CDC)) and the California Breast Cancer Early Detection Program (BCEDP) (funded by the California Tobacco Tax Breast Cancer Fund) to evaluate their breast cancer screening programs by linking program participants to the CCR files for breast cancer status and stage at diagnosis. Results of these linkages will be described in the final report.

The CCR is also collaborating in a study of breast cancer among California's MediCal (the California Medicaid program) population. During calendar year 1993, 15.3% of California's female population between the ages of 30 and 85+ were eligible for MediCal for at least one month. We linked this file of 1,356,484 women against our database as of April 1996 and found 2,354 breast cancers. The linked file is now being analyzed for stage at diagnosis among women on MediCal, and a report of this analysis will be included in the final report.

Objective 6 - Link CCR files with hospital discharge and Medicare files to incorporate insurance status, expected hospital charges, and comorbidities into the CCR database.

Results of our linkages with the Health Care Financing Administration (HCFA) Medicare files and with California Hospital Discharge files were described in our Year 01, Year 02 reports and in the poster presentation at the Department of Defense Breast Cancer Research Program Meeting, November 1-4, 1997 (4) (Appendix V).

Objective 7 - Design and produce a series of confidential and nonconfidential datasets with complete documentation and convenient access for researchers, and produce required reports for the USAMRDC.

Confidential and nonconfidential breast cancer datasets with SEER EOD coding are now available from the CCR to qualified researchers. Follow-up information from the linkages are also available. Two papers using these additional data items are in preparation (5,6) and one paper and one poster have been prepared and presented at national meetings (4,7).

CONCLUSIONS

Work on this project is proceeding. The difficulties encountered in completing EOD coding have been overcome and that portion of the project has been completed. Linkage activities are on schedule. Follow-back to physicians for first course of treatment data has been more resource demanding than originally estimated. The need for computer software to compare treatment information in the CCR records with PDQ recommendations was not anticipated in the original proposal, nor was the necessity for extensive physician follow back including sending staff to physician offices to abstract the additional treatment data. However, gathering additional treatment data for two years of diagnoses will provide a database adequate for studies of breast cancer treatment such as examining characteristics of women who receive "recommended" treatment for their breast cancer compared with women who do not. In addition, it will provide a database

adequate for describing the quality of breast cancer treatment data in the CCR under its current data collection protocol, for describing differences in the characteristics of complete records vs incomplete records, and for accurately assessing required resources for the collection of complete treatment data in California.

REFERENCES

1. National Institutes of Health. Summary Staging Guide - April 1977. U.S. Department of Health, Education, and Welfare, Public Health Service, Bethesda, MD, 1977.
2. Shambaugh EM, Ries LG, Young JL, Kruse MA, Platz CE, Ryan RF, Weiss MA. SEER Extent of Disease, 1988 Codes and Coding Instructions, Second Edition. National Cancer Institute, National Institutes of Health, NIH Publication No. 92-2313, Bethesda, MD, 1992.
3. Beahrs OH, Henson DE, Hutter RVP, Kennedy BJ (Eds.). American Joint Committee on Cancer, Manual for Staging of Cancer - Fourth Edition. J.B. Lippincott, Philadelphia PA, 1992.
4. Wright WE & Allen ME. California Cancer Registry - Enhancement for Breast Cancer Research. Poster Presentation at the Department of Defense Breast Cancer Research Program Meeting. Washington, DC, November 1-4, 1997.
5. Parker JD, Newman J, Gegretsadik T, Kileen M. Managed care and treatment for early stage breast cancer: California Medicare, 1993. MediCal Care (in press).
6. Morris CR. Comparison of SEER Staging and AJCC Staging for Breast Cancer. (Manuscript in preparation.)
7. Allen ME, Wright WE, Perkins CI. Stage at Diagnosis and Treatment among MediCal Breast Cancer Patients in California. Presentation at the Annual Meeting of the North American Association of Central Cancer Registries. Boston, MA, April 2, 1997.

APPENDIX I

CORRECTION RECORD
PROCESSING SPECIFICATIONS

October 21, 1997

The following shaded changes were identified as necessary by Judy Boone, Deborah Bringman and Ben Wormeli during Corrections Processing software testing.

Processing Procedures and Specifications for Applying Correction Records to the Data Base

These procedures and specifications apply only to the backlog of correction files, those that are records version A, B, C or D.

Pre-Processing Procedures:

1. All backlog correction files will be reformatted to the record version D record layout format (C/Net version 2.6). The file creation data will be added to the end of each record in the file and will be used to approximate the date the correction was created since the correction date field may be blank in some records. Since the order the record is written to the file is important for processing it is recommended that each unique record be numbered sequentially. A DOS program is being developed to accomplish this task.
2. Data items in the correction records will be converted to coding procedure 16 either before or during the time the files are processed. For example, the following correction under version C:

<u>Correction Data Item</u>	<u>Correction Value</u>
Surgery Summary	02

would become three corrections:

<u>Correction Data Item</u>	<u>Correction Value</u>
Surgery Summary	00
Surgery Summary-NCD	02
Surgery Summary-Recon	0

Specifications will be written to convert the data items in the correction files.

3. The correction files will be linked to the database to obtain the Region Patient Number and Region Tumor Number for the patient and tumor that applies to each correction record. The Region Patient Number and Region Tumor Number will be inserted back into the correction record.

Processing Procedures:

1. The correction files will be combined into one file and sorted by region patient number, correction creation date, hospital number, region tumor number and correction record number. The file will then be processed. This will allow the records to be processed in patient order by correction date starting with the earliest date. This should alleviate the need to pull abstracts for the same patient more than once. An audit trail should be kept of any corrections automatically applied to the database. This can be

accomplished by indicating the data item that was "changed per correction record" in the Comments or Remarks text fields.

Post-Processing Procedures:

1. Since the corrections may change key linkage variables such as name, date of birth, and social security number, it is recommended that a linkage program be run against the full database after all corrections have been applied to identify any false negative matches.

Specifications for applying the correction value to the data base:

Since the CANDIS, ANEW, and CRIS systems process and store their data differently, these specifications are meant to be general enough to provide guidance for all three systems.

When processing the correction file it is assumed that the value for the correction data item will be compared to its database value/ If the values are the same then the correction is ignored. If the correction value differs from the database value, the correction value is applied to the database according to these specifications.

The specifications for updating the database value with the correction value are indicated by data item. for some data items two levels of specifications are given and are indicated as "Abs" and "Con". The "Abs" specification should be used when only one source (abstract) exists for the patient and/or tumor (depending on the data item). The "Con" specification should be used when multiple source (abstracts) exist for the specific patient or tumor set that matches the correction record.

For some data items the specification states "manual review". For other data items the specification states "list for review" or states "automatically update and list for review". "Manual Review" implies that the patient's abstract(s) must be reviewed before a decision is made to apply the correction. "List for Review" implies that the correction value and the database value should be examined before a decision is made to apply the correction. In most cases the patient's abstract(s) will not need to be examined. "Automatically Update and List for Review" implies that on rare occasions an update of this data item may cause an edit inconsistency in which case the edit inconsistency should be printed for resolution.

Any correction for which the specification states "List for Review" or "Update and List For Review" will produce a listing which will include the case identifiers, correction information, and existing database information. Case identifiers will include the regional patient number, last name, and first name for patient level corrections, as well as regional tumor number and site for tumor level corrections, and hospital number for admission level corrections. Correction information will include the correction creation date, the correction item name, the correction item value, and correction remarks. Database information will include the existing database value for the item being corrected, and in some cases, additional database information which may facilitate the processing of the correction without review of an abstract.

Patient Information
(at the patient level)

Last Name

- Abs: Automatically update
Add the old Last Name to the AKA file as an alias last name if it does not exist
Regenerate NYSIIS-NAME
- Con: If the correction Last Name exists in the AKA file
then ignore the correction
else add old name to AKA,
update and list for review the correction value, the database value and all AKA values
for last name, maiden name, middle name, first name, and patient date of last contact

First Name

- Abs: Automatically update and list for review any first-name/sex inconsistencies
Add the old First Name to the AKA file as an alias first name if it does not exist
- Con: Automatically update if first character of database value = first character of correction
value and characters 2-14 of database value are blank and 2-14 of the correction value
are not blank.
Else if the correction First Name exists in the AKA file
then ignore the correction
Else add old first name in AKA file and
update and list for review the last name, first name, middle name, maiden name and
date of last patient contact.

Middle Name

- Abs: Automatically update and add old middle name to AKA first name.
- Con: Automatically update if database value is blank and correction value is not blank.
Automatically update if first character of database value = first character of correction
value and characters 2-14 of database value are blank and 2-14 of the correction value
are not blank.
Ignore if the correction value is blank and the database value is not blank
Ignore if the correction matches first Name or AKA First Name
Else add to AKA file as an AKA First Name and update and list for review the last
name, middle name, maiden name, and date of last patient contact.

Alias Last Name

Add to the AKA file as an alias last name if it does not exist in the AKA file as a last
name or maiden name and automatically update and list for review at a later time to
ensure it is not a first name or a last name suffix.
(List at a later time = not done by corrections programs)
Generate NYSIIS-NAME

Alias First Name

Ignore if equal to First Name
Else add to AKA file as a First Name if it does not exist in the AKA file.

Maiden Name

Ignore if equal to Last Name
Else add to the AKA file as a maiden name
Generate NYSIIS-NAME

Name Suffix

Abs: Automatically update

Con: Automatically update if database value is blank
Else Update and List For Review

Social Security Number

Abs: Automatically update if database value is blank or 9's and correction value is not blank or 9's
Else ignore if Follow-up Source on database = 26 or 56 (Information from death clearance)
and date of last patient contact or death is prior to 1996
Else automatically update

Con: Automatically update if database value is blank or 9's and correction value is not blank or 9's
Else ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996
Else ignore if correction value = blank or 9's
Else ignore if correction suffix = AD@
Else automatically update and list for review
~~Note: Any time social security number is changed, program will list for review any existing patient records with the same social security number~~

Social Security Number Suffix

If Social Security Number is updated
Automatically update
Else
If Follow-up Source on database not = 26 or 56 and
Date of Last Patient Contact or Death is prior to 1996
Ignore
Else
If correction comes in without correction to Social Security Number and database
social security number is not blank or 9's
Automatically update

Birthplace

Abs: Automatically update if
1. database value is 999 OR
2. database value is 000 and correction value is 001-099

Else ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996.

Con: Automatically update if
1. database value is 999 OR
2. database value is 000 and correction value is 001-099

Else ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996

Else automatically update and list for review any birthplace/ race or birthplace/spanish origin inconsistencies.

Date of Birth

Abs: Ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996
Else automatically update and list for review dob, site, marital status, date-dx, histology,
and any edits or date conflicts
Recalculate Age at Diagnosis and Age Group and list for review any age/marital status
or age/site inconsistencies

Cons: (Same as Abs.)
Ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996
Else automatically update and list for review dob, site, marital status, date dx, histology,
and any edits or date conflicts
Recalculate Age at Diagnosis and Age Group and list for review any age/marital status
or age/site inconsistencies

Race

Ignore if Follow-up Source on database = 26 or 56
and date of last patient contact or death is prior to 1996
Else automatically update if
1. database value is 99 and correction value is 00-98 OR
2. database value is 96 and correction value is 04-06 or 08-14 OR
3. database value is 97 and correction value is 07 or 20-32
and list for review birthplace, race, spanish origin and any Race/Birthplace
inconsistencies
Else manual review

Spanish-Origin

List for review last name, maiden name, birthplace, race

Sex

Automatically update if database value is 9 and correction value is not 9 and list for
review any site/sex or first-name/sex inconsistencies
Else manual review

Patient Information
(at the upper level)

Marital Status

Abs: Automatically update and list for review any age/marital status inconsistencies

Con: Automatically update if the correction record is from a class 0-2 case and
1. the database value is 9 and the correction value is not 9 OR
2. the database value is 1 and the correction value is 2-5
List for review any age/marital status inconsistencies
Else ignore

Occupation - Text

Abs: Automatically update
Reset OCCUP-80 and OCCUP-90 to 9999

Con: Automatically update if the database value is blank or "NR" or "Retired" and reset
OCCUP-80 and OCCUP-90 to 9999
Else list for review at a later time (later time = not done by corrections program)

Industry - Text

Abs: Automatically update
Reset INDUS-80 and INDUS-90 to 9999

Con: Automatically update if the database value is blank or "NR" and reset INDUS-80 and
INDUS-90 to 9999
Else list for review at a later time (later time = not done by corrections program)

Religion

Abs: Automatically update

Con: Automatically update if
1. the database value is 00 or 99 and correction value is not 00 or 99 OR
2. the database value is 20 and the correction value is 10-70
Else ignore

DX Address

Run address standardization on correction value before comparing it to the database
value,
If values differ then list for review dx address, dxcity, dxstate, dxzip, dxcounty.

DX City

Run city variant conversion on correction value before comparing it to the database
value,
If values differ then list for review dx address, dxcity, dxstate, dxzip, dxcounty.

DX State

List for review dx address, dxcity, dxstate, dxzip, dxcounty.

DX Zip

List for review dx address, dxcity, dxstate, dxzip, dxcounty.

DX County

List for review dx address, dxcity, dxstate, dxzip, dxcounty.

Tumor Information
(at the tumor level)

Sequence Number

Ignore if database value = '00' and correction value = '01'
Else Abs: update and list for review (include correction remarks).
Con: list for review.

Date of Diagnosis

Manual review

Site - ICDO2

Manual review

Site - ICDO1

Manual review

Laterality

Ignore if laterality not required for site
Else Manual review

Histology - Type

Manual review

Histology - Behavior

Manual review

Histology - Differentiation

Automatically update if correction hist type = database hist type
and database value = 9 and correction value = 1-4
Else Manual review

Summary Stage

If DATEDX Year < 1994 then manual review
Else ignore

Tumor Size

Ignore if DATEDX year prior to 1994 or unknown
Else manual review

Direct Extension

Ignore if DATEDX year prior to 1994 or unknown
Else manual review

Direct Extension - Pathology

Ignore if DATEDX year prior to 1995 or unknown
Else manual review

Lymph Node Summary

Ignore if DATEDX year prior to 1994 or unknown
Else manual review

Nodes Positive

Ignore if DATEDX year prior to 1994 or unknown
Else manual review

Nodes Examined

Ignore if DATEDX year prior to 1994 or unknown
Else manual review

Pediatric Stage

Automatically update.

Pediatric Stage Coder

Automatically update.

Pediatric Stage System

Automatically update.

Residual Tumor

Abs: Automatically update

Con: Automatically update if the admission record that matches the correction record has the more definitive value for surghosp than all other admission records for that tumor
Else Ignore

Diagnostic Confirmation

Abs: Automatically update and list for review any diagnostic confirmation interfield edit errors

Con: Automatically update if correction value < database value and list for review any diagnostic confirmation interfield edit errors

Type of Reporting Source

- Abs: Automatically update and list for review any typerep interfield edit errors
- Con: Automatically update and list for review any typerep interfield edit errors if appropriate based hierarchy 1,4,5,3

Tumor Markers

- Abs: Automatically update at the admission and tumor level
- Con: automatically update at the admission level and apply consolidation rules at the tumor level

Treatment Information (at the tumor and admission level)

Surgery - consider the following items as a group and apply all or none accordingly. (Though not all items will be found in the backlog of correction records, because the correction records are converted to coding procedure 16 before applying, it is most likely there will be converted correction values for the new data items.)

Date of Surgery

Date of Surgery - Non-Cancer Directed

Surgery Summary

Surgery Summary - Non-Cancer Directed

Surgery summary - Reconstructive

Surgical Approach

Surgery at this Hospital

Surgery at this Hospital - Non-Cancer Directed

Surgery at this Hospital - Reconstructive

Reason for No Surgery

1. If any of the data items SURG-APPROACH, SURG-HOSP, SURG-HOSP-NCD, SURG-HOSP-RECON are the only corrections in the group (ie. the correction is changing the type of surgery done at the hospital) then:

Automatically update if values are consistent with treatment summary values (ie. no surgery/surghosp interfield edit errors)
Else manual review

2. If DATE-SURG and/or DATE-SURG-NCD are the only corrections in the group (ie. the corrections only changing the date of surgery) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR
2. the correction date is with four months of Diagnosis Date and the correction produces no surgery or date interfield edit errors.
Regenerate Date of Therapy if needed.
Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal and the correction produces no surgery or date interfield edit errors.
Regenerate Date of Therapy and Radiation Sequence if needed.
Else manual review

3. If the corrections in the group contain DATE-SURG AND surg-sum (and possibly other surgery variables) and the correction DATE-SURG NOT = 0'S AND THE DATABASE date-surg = 0's (ie. the correction is adding surgical treatment):

Automatically update if

1. the correction DATE-SURG is within four months of Date-Dx OR
 2. the Date of Therapy = 0's
- and all surgery and date fields are consistent (no interfield edit errors).
Regenerate Date of Therapy and Radiation Sequence if needed.
Add "Surgery added per correction record" to Text-Surg field. Do not overwrite any text already in the field.
Else manual review.

4. If the corrections in the group contain DATE-SURG AND SURG-SUM (and possibly other surgery variables) and the correction DATE_SURG = 0's and the database DATE-SURG not = 0's (ie. the corrections deleting surgical treatment):

Manual review

5. For any other combinations of variables in the group (ie. the correction is changing the type of surgical treatment that was given):

Manual review

Radiation - consider the following items as a group and apply all or none accordingly.

Date of Radiation

Radiation summary

Radiation to CNS Summary

Radiation at this Hospital

Radiation to CNS at this Hospital

Reason for No Radiation

Radiation/Surgery Sequence

1. If RAD-HOSP and/or RAD CNS-HOSP are the only corrections in the group (ie. the correction is changing the type of radiation done at the hospital) then:

Automatically update if values are consistent with treatment summary values (ie. no radsum/radhosp interfield edit errors)
Else manual review

2. If DATE-RAD is the only correction in the group (ie. the correction is only changing the date of radiation) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR
 2. the correction date is within four months of Diagnosis Date and the correction produces no radiation or date interfield edit errors.
- Regenerate Date of Therapy and Radiation sequence if needed.
Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal and the correction produces no radiation or date interfield edit errors.
- Regenerate Date of Therapy and Radiation Sequence if needed.
Else manual review

3. If the corrections in the group contain DATE-RAD and RAD-SUM (and possibly other radiation variables) and the correction DATE-RAD not = 0's and the database DATE-RAD = 0's (ie. the correction is adding radiation treatment):

Automatically update if

1. the correction DATE-RAD is within four months of Date-Dx OR
 2. the Date of Therapy = 0's
- and all radiation and date fields are consistent (no interfield edit errors).
Regenerate Date of Therapy and Radiation Sequence if needed.
Add "Radiation added per correction record" to Text-Radiation field. Do not overwrite any text already in the field.
Else manual review

4. If the corrections in the group contain DATE-RAD and RD-SUM (and possible other radiation variables) and the correction DATE-RAD = 0's and the database DATE-RAD not= 0's (ie. the correction is deleting radiation treatment):

Manual review

5. For any other combinations of variables in the group (ie. the corrections changing the type of radiation treatment that was given):

Manual review

Chemotherapy - consider the following items as a group and apply all or none accordingly.

Date of Chemotherapy

Chemotherapy Summary

Chemotherapy at this Hospital

Reason for No Chemotherapy

1. If CHEMO-HOSP is the only correction in the group (ie. the correction is changing the type of chemotherapy given at the hospital) then:

Automatically update if the value is consistent with treatment summary values (ie. no chemosum/chemohosp interfield edit errors)
Else manual review

2. If DATE-CHEMO is the only correction in the group (ie. the correction is only changing the date of chemotherapy) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR
 2. the correction date is with four months of Diagnosis Date and the correction produces no chemotherapy or date interfield edit errors.
- Regenerate Date of Therapy if needed.
Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and know parts of the correction date and the database date are equal and the correction produces no chemotherapy or date interfield edit errors.
- Regenerate Date of Therapy if needed.
Else manual review

3. If the corrections in the group contain DATE-CHEMO and CHEMO-SUM (and possible other chemotherapy variables) and the correction DATE-CHEMO not = 0's and the database DATE-CHEMO = 0's (ie. the correction is adding chemotherapy):

Automatically update if

1. the correction DATE-CHEMO is within four months of Date-Dx OR
 2. the Date of Therapy = 0's
- and all chemotherapy and date fields are consistent(no interfield edit errors).
Regenerate Date of Therapy if needed.
Add "Chemotherapy added per correction record" to Text-Chemotherapy field.
Do not overwrite any text already in the field.
Else manual review.

4. If the corrections in the group contain DATE-CHEMO and CHEMO-SUM (and possible other chemotherapy variable) and the correction DATE-CHEMO = 0's and the database DATE-CHEMO not = 0's (ie. the correction is deleting chemotherapy):

Manual review

5. For any other combinations of variables int he group (ie. the correction is changing the type of chemotherapy that was given):

Manual review

Hormone Therapy - consider the following items as a group and apply all or none accordingly.

Date of Hormone Therapy

Hormone Therapy Summary

Hormone Therapy at this Hospital

Reason for No Hormone Therapy

1. if HORM-HOSP is the only correction is the group (ie. the correction is changing the type of hormone therapy given at the hospital) then:

Automatically update if the value is consistent with treatment summary values (ie. no hormsum/hormhosp interfield edit errors)

Else manual review

2. if DATE-HORM is the only correction is the group (ie. the correction is only changing the date of hormone therapy) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR

2. the correction date is within four months of Diagnosis Date

and the correction produces no hormone therapy or date interfield edit errors.

Regenerate Date of Therapy if needed.

Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal

and the correction produces no hormone therapy or date interfield edit errors.

Regenerate Date of Therapy if needed.

Else manual review

3. If the corrections in the group contain DATE-HORM and HORM-SUM (and Possible other hormone therapy variables) and the correction DATE-HORM not = 0's and the database DATE-HORM = 0's (ie. the correction is adding hormone therapy)"

Automatically update if

1. the correction DATE-HORM is with four months of Date-Dx OR

2. the Date of Therapy = 0's

and all hormone therapy an date fields are consistent (no interfield edit errors).

Regenerate Date of Therapy if Needed.

Add "Hormone therapy added per correction record" to Text-Hormone field

Do not overwrite any text already in the field.

Else manual review

4. If the corrections in the group contain DATE-HORM and HORM-SUM (and possible other hormone therapy variables) and the correction DATE-HORM = 0's and the database DATE-HORM not = 0's (ie. the correction is deleting hormone therapy):

Manual review

5. For any other combinations of variables in the group (ie. the correction is changing the type of hormone therapy that was given):

Manual review

Immunotherapy - consider the following as a group and apply all or none accordingly.

Date of Immunotherapy
Immunotherapy Summary
Immunotherapy at this Hospital

1. If IMMUNO-HOSP is the only correction in the group (ie. the correction is changing the type of immunotherapy given at the hospital) then:

Automatically update if the value is consistent with treatment summary values (ie. no immunosum/immunohosp interfield edit errors)
Else manual review

2. if DATE-IMMUNO is the only correction is the group (ie. the correction is only changing the date of immunotherapy) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR
 2. the correction date is within four months of Diagnosis Date
- and the correction produces no immunotherapy or date interfield edit errors.
Regenerate Date of Therapy if needed.
Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal
- and the correction produces no immunotherapy or date interfield edit errors.
Regenerate Date of Therapy if needed.
Else manual review

3. If the corrections in the group contain DATE-IMMUNO and IMMUNO-SUM (and possibly IMMUNO-HOSP) and the correction DATE-IMMUNO not = 0's and the database DATE-IMMUNO = 0's (ie. the correction is adding immunotherapy):

Automatically update if

1. the correction DATE-IMMUNO is with four months of Date-Dx

OR

2. the Date of Therapy = 0's
- and all immunotherapy an date fields are consistent (no interfield edit errors).
Regenerate Date of Therapy if needed.
Add "Immunotherapy added per correction record" to Text-Immunotherapy field.
Do not overwrite any text already in the field.
Else manual review

4. If the corrections in the group contain DATE-IMMUNO and IMMUNO-SUM (and possibly IMMUNO-HOSP) and the correction DATE-IMMUNO = 0's and the database DATE-IMMUNO not = 0's (ie. the correction is deleting immunotherapy):

Manual review

5. For any other combinations of variables in the group (ie. the correction is changing the type of immunotherapy that was given):

Manual review

Other Therapy - consider the following items as a group and apply all or none accordingly.

Date of Other Therapy

Other Therapy Summary

Other Therapy at this Hospital

1. If OTHER-HOSP is the only correction in the group (ie. the correction is changing the type of other therapy given at the hospital) then:

Automatically update if the value is consistent with treatment summary values (ie. no othersum/otherhosp interfield edit errors)
Else manual review

2. if DATE-OTHER is the only correction in the group (ie. the correction is only changing the date of other therapy) then:

Abs: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal OR
 2. the correction date is within four months of Diagnosis Date and the correction produces no other therapy or date interfield edit errors.
- Regenerate Date of Therapy if needed.
Else manual review

Con: Automatically update if

1. the correction date value is replacing 9's in the database date value and known parts of the correction date and the database date are equal and the correction produces no other therapy or date interfield edit errors.
- Regenerate Date of Therapy if needed.
Else manual review

3. If the corrections in the group contain DATE-OTHER and OTHER-SUM (and possibly OTHER-HOSP) and the correction DATE-OTHER not = 0's and the database DATE-OTHER = 0's (ie. the correction is adding other therapy):

Automatically update if

1. the correction DATE-OTHER is within four months of Date-Dx OR
 2. the Date of Therapy = 0's
- and all other therapy and date fields are consistent (no interfield edit errors).
Regenerate Date of Therapy if needed.
Add "Other therapy added per correction record" to Text-Other therapy field.
Do not overwrite any text already in the field.
Else manual review

4. If the corrections in the group contain DATE-OTHER and OTHER-SUM (and possibly OTHER-HOSP) and the correction DATE-OTHER = 0's and the database DATE-OTHER not = 0's (ie. the correction is deleting other therapy):

Manual review

5. For any other combinations of variables in the group (ie. the correction is changing the type of other therapy that was given):

Manual review

Admission Information
(at the admission level)

Hospital Number

Manual review

Accession Number

Automatically update

Year First Seen

Automatically update

Medical Record Number

Automatically update

Date of First Admission

Automatically update and list for review any date admission interfield edit errors

Date of Inpatient Admission

Automatically update

Date of Inpatient Discharge

If correction value is not 00000000 and type of admission is 2, 3, 4, or 7 then change correction value to 00000000.

Automatically update and list for review any date discharge interfield edit errors

Physician - Attending

Automatically update if correction value is a valid MD code (no edit errors) and move old value to Physician-Other field else ignore

Physician - Referring

Automatically update if correction value is a valid MD code (no edit errors) and move old value to Physician-Other field else ignore

Physician - Surgeon

Automatically update if correction value is a valid MD code (no edit errors) and move old value to Physician-Other field else ignore

Physician - Medical Oncologist

Automatically update if correction value is a valid MD code (no edit errors)
Else ignore

Physician - Radiation Oncologist

Automatically update if correction value is a valid MD code (no edit errors)
Else ignore

Physician - Other

Automatically update if correction value is a valid MD code (no edit errors)
Else ignore

Class of Case

Automatically update and list for review any class interfield edit errors

Hospital From

Automatically update and list for review class, hospital number, and any hospfrom/datedx/dateadm inconsistencies

Hospital To

Automatically update and list for review class, and hospital number, ~~and any hospfrom/datedx/dateadm inconsistencies.~~

Casefinding Source

Automatically update

Payment Source

Automatically update

Payment Source - Text

Automatically update if Payment Source Text is blank or Payment Source is updated

Regional Data

Ignore (At the discretion of the region)

Pathology Report Number - Biopsy/FNA

Automatically update

Pathology Report Number - Surgery

Automatically update

ACOS Information
(at the admission level)

TNM Coder - Clinical

Automatically update

TNM Coder - Path

Automatically update

TNM Edition

Automatically update

TNM T Code - Clinical
- Path

Automatically update

TNM N Code - Clinical
- Path

Automatically update

TNM M Code - Clinical
- Path

Automatically update

TNM Staging Basis

Automatically update

TNM Stage - Clinical
- Path

Automatically update

APPENDIX II

FEMALE BREAST CANCER TREATMENT GROUPS

FROM THE

NATIONAL CANCER INSTITUTE'S

PHYSICIAN DATA QUERY SYSTEM

March 6, 1997

FEMALE BREAST CANCER STAGE GROUPS

Site = C500-509 (740-749); Sex = 2

Histology = 8000-8231, 8244-8573

(excludes carcinoids, sarcomas, lymphomas)

Dxdate > = 1995

74 Group	DEFINITION TNM = 4th Edition	CONDITIONS	PDQ CRITERIA FOR APPROPRIATE RX
1	STAGE 0 Tis NO MO Non-lobular, in situ	Hist = XXXX2X Hist = 8520X	L, R, %, % M, %, %, %
2	STAGE 0 Tis NO MO Lobular in situ/lobular neoplasia	Hist = 85202X	L, %, %, % L, %, %, H M, %, %, %
3	STAGE I T1 NO MO ERA Negative	EOD col 1-3 < = 020; 997; 999 EOD col 4-5 = 05, 10, 99 EOD col 6 = 0, 9 ERA = 2	L, R, C, % M, %, C, %
4	STAGE I T1 NO MO ERA Positive	EOD col 1-3 < = 020; 997; 999 EOD col 4-5 = 05, 10, 99 EOD col 6 = 0, 9 ERA = 1	L, R, %, H M, %, %, H
5	STAGE I T1 NO MO ERA Unknown	EOD col 1-3 < = 020; 997; 999 EOD col 4-5 = 05, 10, 99 EOD col 6 = 0, 9 ERA = 0, 3, 8, 9	L, R, C, H M, %, C, H
6	STAGE II T0, T1, T2 N1 MO Positive nodes Premenopausal	EOD col 1-3 < = 050 EOD col 4-5 = 10, 99 EOD col 6 = 1-4, 6 Age < 50	L, R, C, % M, %, C, %
7	STAGE II T0, T1, T2 N1 MO Positive nodes Postmenopausal, ERA = positive	EOD col 1-3 < = 050 EOD col 4-5 = 10, 99 EOD col 6 = 1-4, 6 Age > = 50, ERA = 1	L, R, %, H M, %, %, H
8	STAGE II T2, T3 NO MO Negative nodes ERA negative	EOD col 1-3 > 020 and < 997, 998 EOD col 4-5 = 10, 99 EOD col 6 = 0, 9 ERA = 2	L, R, C, % M, %, C, %
9	STAGE II T2, T3 NO MO Negative nodes ERA Positive	EOD col 1-3 > 020 and < 997, 998 EOD col 4-5 = 10, 99 EOD col 6 = 0, 9 ERA = 1	L, R, C, H M, %, C, H
10	STAGE II T2, T3, NO MO Negative nodes ERA Unknown	EOD col 1-3 > 020 and < 997, 998 EOD col 4-5 = 10, 99 EOD col 6 = 0, 9 ERA = 0, 3, 8, 9	L, R, C, H M, %, C, H
11	STAGE III A T0, T1, T2 N2 MO	EOD col 1-3 < = 050 EOD col 4-5 = 10, 99 EOD col 6 = 5	M, R, C, %

FEMALE BREAST CANCER STAGE GROUPS

Site = C500-509 (740-749); Sex = 2

Histology = 8000-8231, 8244-8573

(excludes carcinoids, sarcomas, lymphomas)

Dxdate > = 1995

74 Group	DEFINITION TNM = 4th Edition	CONDITIONS	PDQ CRITERIA FOR APPROPRIATE RX
12	STAGE III A T3 N1, N2 MO	EOD col 1-3 >050 and <997, 998-999 EOD col 4-5 = 10, 99 EOD col 6 = 1-6	M, R, C, %
13	STAGE III B T4 any N MO	EOD col 4-5 = 20-70 EOD col 6 = 0-6, 9	L, R, C, %
14	STAGE III B any T N3 MO	EOD col 4-5 = 10-70, 99 EOD col 6 = 7	L, %, C, %
15	STAGE III B any T any N MO ERA and PRA positive	EOD col 4-5 = 10-70, 90 EOD col 6 = 1-7 ERA = 1 and PRA = 1	% %, %, H
16	STAGE IV Any T and N M1 ERA Positive and PRA Positive No visceral disease	EOD col 4-5 = 10-80, 99 EOD col 6 = 8 ERA = 1 and PRA = 1	%, R, %, % M, %, %, % %, %, %, M
17		_____or_____ EOD col 4-5 = 80 EOD col 6 = 0-7, 9 ERA = 1 and PRA = 1	%, R, %, % M, %, %, % %, %, %, M
18	STAGE IV Any T anyN M1 ERA Negative and PRA Negative Visceral disease present	EOD col 4-5 = 10-80, 99 EOD col 6 = 8 ERA = 2 and PRA = 2 EOD col 4-5 = 85 EOD col 6 = 0-9	%, R, %, % M, %, %, % %, %, C, %
19	STAGE IV Any T any N M1 ERA Unknown and PRA Unknown	EOD col 4-5 = 80-85 EOD col 6 = 0-9 ERA = 0, 3, 8, 9 and PRA = 0, 3, 8, 9	%, R, %, % M, %, %, % %, %, %, M
20		_____or_____ EOD col 4-5 = 10-70, 99 EOD col 6 = 8 ERA = 0, 3, 8, 9 and PRA = 0, 3, 8, 9	%, R, %, % M, %, %, % %, %, %, M
21	STAGE Unknown	EOD col 4-5 = 99 EOD col 6 = 0, 9	%, R, %, % M, %, %, % %, %, %, M

Notes:

TNM staging information is per the AJCC Manual for Staging of Cancer, Fourth Edition. Treatment recommendations are per the PDQ Information for Health Care Professionals, Breast Cancer, modified 3/97. SEER Program Code Manual, Revised June 1992, page 98: Estrogen receptor field (Tumor marker 1; Section IV, Field 07.A) is coded for breast cases diagnosed on or after January 1, 1990. For all cases diagnosed before January 1, 1990, this field is coded to 9.

1% = can have any value.

APPENDIX III

CAPTURE OF MISSING FIRST COURSE OF THERAPY
IN BREAST CANCER PATIENTS

M. OEHRLI

DAISER PERMANENTE - NORTHERN CALIFORNIA

September 26, 1997

KAISER PERMANENTE-NORTHERN CALIFORNIA
REGIONAL CANCER REGISTRY

cc: Kim
Helen
Craig
9/30/97

MEMORANDUM

TO: LILIA O'CONNOR
FROM: MICHAEL OEHRLI
DATE: SEPTEMBER 26, 1997
SUBJECT: CAPTURE OF MISSING FIRST COURSE OF THERAPY IN CANCER PATIENTS
CC: ROBERT HIATT, MD, PHD
LEO HURLEY

Enclosed please find the final summary of our project and an accompanying diskette. It was a pleasure to work on this project and I hope you find that is answers as many questions for you as it has for us here at Kaiser Permanente.

In brief summary, we found that the AOMS system which tracks radiation therapy authorizations is highly effective both in terms of capture of information and enhancement of registry data. On the contrary, the OSCR system is far from complete in its capture of chemotherapy information, but it can also serve to enhance registry data.

Please give me a call at (510) 450-2087 if you have any questions.

NCCCSTDY.MEM
MDO/09-97

KAISER PERMANENTE-NORTHERN CALIFORNIA
REGIONAL CANCER REGISTRY

CAPTURE OF MISSING FIRST COURSE OF THERAPY DATA
IN CANCER PATIENTS DIAGNOSED 1994-1995

PURPOSE: To evaluate the completeness of reported cancer data regarding radiation and chemotherapy administered as the first course of therapy.

BACKGROUND:

Northern California Cancer Center (NCCC, Regions 1&8) has provided funding to evaluate the completeness of cancer data among Kaiser Permanente's Northern California facilities located in these regions. This study was completed on September 26, 1997 by Michael Oehrli and Leo Hurley under the supervision of Dr. Robert A. Hiatt.

Three files have been submitted for matching with information contained in mainframe systems. A fourth file has been generated by the Regional Cancer Registry (RCR) consisting of all diagnoses reported 1994-1996 to use for comparison. The NCCC files are subsets of the complete RCR file. Most patients with multiple records in the NCCC files were diagnosed with simultaneous bilateral breast cancer.

<u>FILE</u>	<u>SOURCE</u>	<u>DX</u>	<u>N=CASES</u>	<u>N=ERRORS/ MULTRECS</u>	<u>N=PATIENTS</u>
PDQ-1994	NCCC-Breast	94	1,053	14	1,039
PDQ-1995	NCCC-Breast	95	553	7	546
POC-1995	NCCC-All Sites	95	265	0	265
RCR-9496	RCR-All Sites	94-96	23,845	1,273	22,572

The files were first verified for key identifier information required to successfully match the data to mainframe systems. Errors in medical record number (N=28), accession number (N=1), diagnosis date (N=1), and regional patient identification number (N=1) are identified in the files sent back to NCCC (*mrg.asc) with an explanation provided in the NOTES field.

PROCESS:

The matching process consisted of the following steps:

1. Preparation and cleanup of study files.
2. Extraction of all radiation therapy authorizations from mainframe.
3. Extraction of all administrations of chemotherapy from mainframe.
4. Preparation of a Regional Cancer Registry file containing existing information relating to first course of therapy for comparison.
5. Merging of the above four files using a common linkage.
6. Writing programs to calculate time from diagnosis to start of treatment and determination of first course vs. subsequent therapy.

Once cleanup of the study files was complete, all patients in the AOMS (Authorized Outside Medical Services) system assigned an authorization code for payment of radiation therapy services were extracted for the time period 01/94 through 06/97. A total of 18,597 records were extracted. Each record includes an authorization number, issue month/day/year of the authorization, and the name of the vendor or physician who performed the services. The issue date is the only information available to indicate the start of radiation therapy, however, it is not a guarantee that radiation was actually carried out in every case. Occasionally, contraindications may prohibit radiation therapy from starting and this cannot be determined from this file. No information is available regarding dosages or methods of application.

All patients in the OSCR (Outpatient Services Clinical Record) system with a record of chemotherapy administration were extracted for the time period 01/94 through 12/96. A total of 32,150 records were extracted. Each record includes the first date the service was provided, and the facility where it was provided. Particular care must be exercised with this file as the system was in the process of rollout in 1994 and is incomplete for that year. Data from 1997 is not provided due to problems early this year with completeness. No information is available regarding drugs utilized.

The four files (NCCC Study, AOMS, OSCR, Regional Cancer Registry) were linked on medical record number and sequence number (in the case of multiple primaries). The following is a record layout and explanation of variables contained in the linked file:

NCCC STUDY FILES

1-12	MEDRECNO*	KP Medical Record Number
13-20	REGPATNO	CCR Regional Patient Identification Number
21-26	HOSPNO	Reporting Hospital
27-35	ACCNO	Hospital-Specific Accession Number
36-39	SEQNO**	Number of Tumors per Patient
40-47	DATEDX	Date of Diagnosis
48-51	SITENEW	ICD-O-2 Primary Site Code
52	CODE8	Y=Treatment Recommended, Unknown if done
53	PDQ	Y=Treatment Recommended, Physician Data Query Standard.

AOMS FILE

54-59	RTDATE	First Date of Radiation Therapy Authorization
60-69	AUTHORIZE	Authorization Number
70-99	VENDOR	Vendor Performing Radiation Services

OSCR FILE

100-107	CHEMDATE	First Date of Chemotherapy Administration
108-110	FACILITY	Facility Where Chemotherapy Administered

REGIONAL CANCER REGISTRY FILE

111-118	RXDATER	Date of Radiation Performed
119	RADSUM	Type of Radiation Performed
120	RSNORAD	Reason No Radiation Performed
121-128	RXDATEC	Date of Chemotherapy Performed
129	CHEMOSUM	Type of Chemotherapy Performed
130	RSNOCHEM	Reason No Chemotherapy Performed

NEW VARIABLES GENERATED

131-133	FSTCSERT	Yes=First Course Radiation Performed***
134-136	SUBSEQRT	Yes=Subsequent/Recurrent Radiation Performed****
137-144	RTDOCTOR	Vendor Text Converted to State Doctor Code
145-150	RTVENDOR	Vendor Text Converted to State Hospital Code
151-153	FSTCSECH	Yes=First Course Chemotherapy Performed***
154-156	SUBSEQCH	Yes=Subsequent/Recurrent Radiation Performed****
157-176	NOTES	Description of Changes Made to Original Record

EXPLANATION

* Key Linkage Variable
 ** Secondary Linkage Variable
 *** If (Treatment Date-Date of Diagnosis) < One Year
 **** If (Treatment Date-Date of Diagnosis) >= One Year

RESULTS/CONCLUSIONS:

With regard to the study files, the percentage of cases missing first course radiation therapy data ranged from 13% to 21%. The percentage of cases missing chemotherapy data ranged from 1% to 5%. In comparison, 9% of all Regional Cancer Registry cases were missing radiation data. Chemotherapy data was missing in only 3% of the cases, however, the OSCR database is highly incomplete. The extent of missing chemotherapy information in registry data appears minimal due to the fact that this service is usually performed in-house and most of the data would be collected through chart review.

RADIATION THERAPY	PDQ-1994	PDQ-1995	POC-1995	<u>Total</u>
N=PATIENTS	1039	546	265	1850
N=PATIENTS W/EXISTING RT INFO	315	154	73	
N=PATIENTS W/AOMS 1ST COURSE RT	450	269	127	
N=PATIENTS MISSING 1ST COURSE RT	135	115	55	
N=PATIENTS W/NEW RT DATA	138 (13%)	117 (21%)	54 (20%)	309 (16.7%)
CHEMOTHERAPY	PDQ-1994	PDQ-1995	POC-1995	
N=PATIENTS	1039	546	265	
N=PATIENTS W/EXISTING CHEMO INFO	260	116	78	
N=PATIENTS W/OSCR 1ST COURSE CHEMO	127	99	63	
N=PATIENTS MISSING 1ST COURSE CHEMO	133	17	15	
N=PATIENTS W/NEW CHEMO DATA	15 (1%)	23 (4%)	14 (5%)	52 (2.8%)

Looking below at code 8 cases only (treatment recommended, but unknown if given), the percentage of these cases where radiation data were found ranged from 48% to 89%. The percentage of cases where new chemotherapy information was found averaged 38% but the numbers are too small to make a conclusion.

RADIATION THERAPY	PDQ-1994	PDQ-1995	POC-1995
N=CODE8 CASES	35	36	265
N=CODE8 CASES W/AOMS 1ST COURSE RT	30 (86%)	32 (89%)	127 (48%)
CHEMOTHERAPY	PDQ-1994	PDQ-1995	POC-1995
N=CODE8 CASES	4	3	1
N=CODE8 CASES W/OSCR 1ST COURSE CHEMO	1 (25%)	1 (33%)	1 (100%)

A quality check on the completeness of the AOMS database revealed that only five cases out of the 542 cases known to have received radiation were not found. Further evaluation of these five cases confirmed their treatment and the reason they were not found in AOMS could not be determined. In comparison to the entire RCR database; of the 344 cases where radiation treatment was confirmed but no record existed in AOMS; 119/344 (35%) of the patients had treatment for thyroid cancer where radioisotope therapy is provided by nuclear medicine departments internally, 133/344 (39%) of the patients had multiple tumors and then AOMS data was linked to both records, and 92/344 (26%) had no AOMS record for reasons unable to be determined.

The OSCR database was found to be highly incomplete, although, after initial rollout in 1994, capture of chemotherapy information increased to 66% in 1995.

AOMS DATABASE	PDQ-1994	PDQ-1995	POC-1995
N=PATIENTS W/EXISTING RT INFO	315	154	73
N=PATIENTS W/RT FOUND IN AOMS	312 (99%)	152 (99%)	73 (100%)
OSCR DATABASE	PDQ-1994	PDQ-1995	POC-1995
N=PATIENTS W/EXISTING CHEMO INFO	260	116	78
N=PATIENTS W/CHEMO FOUND IN OSCR	112 (43%)	76 (66%)	49 (63%)

The following table is provided for comparison and summarizes all patients diagnosed 1994-1996 in the Regional Cancer Registry whose reason for no radiation or chemotherapy treatment is noted. A significant amount of new radiation therapy data has been obtained, however, due to the incomplete nature of the OSCR database, a large number of chemotherapy data is missing. As expected, cancer registry data obtained through chart review appears more complete than OSCR for treatment occurring in-house. On the contrary, a significant number of radiation therapy patients have been found through AOMS due to treatment external to Kaiser Permanente. In both cases, new information has been found.

REASON RT OR CHEMO NOT PERFORMED	EXISTING RADIATION DATA	AOMS RT FIRST COURSE	AOMS RT SUBSEQUENT	EXISTING CHEMO DATA	OSCR CHEMO FIRST COURSE	OSCR CHEMO SUBSEQUENT
Code 0 Performed	4,801	4,476	11	4,597	2,067	75
Code 1 Not Recommended	5,247	443	39	5,238	230	12
Code 2 Contra- indicated	37	3	0	30	0	0
Code 6 Unknown Reason for No Treatment	11,604	1,122	311	12,096	422	186
Code 7 Patient Refused	300	81	6	473	20	7
Code 8 Recommended Unknown if Given	565	478	5	122	33	3
Code 9 Unknown	17	11	0	16	4	1

RECOMMENDATIONS:

Due to the large percentage of cases missing radiation therapy data, any studies evaluating the first course of therapy should match their cohort with additional data from AOMS to complete first course of therapy. AOMS data will be downloaded on a periodic basis and retained at the Regional Cancer Registry for matching with special studies and other uses upon request.

While the OSCR database is not as effective in capturing chemotherapy data as is chart review by cancer registrars, a few new cases can still be found. OSCR data will be downloaded and retained by the Regional Cancer Registry as well.

For purposes of this study, the complete matched files (*mrg.asc) have been exported in flat, ASCII text format and sent to NCCC for further processing. Record layouts (*.dir) are provided on diskette and are the same as noted on page 2 of this report.

APPENDIX IV

COVER LETTER
AND
PHYSICIAN QUESTIONNAIRE

BREAST CANCER TREATMENT FOLLOWBACK STUDY

NORTHERN CALIFORNIA CANCER CENTER

1997

May 14, 1997

Dear Dr:

Re: Treatment for 1995 Breast Cancer Cases

With the great interest in breast cancer incidence and outcomes in California, the California Cancer Registry (CCR) is making a special effort to capture all treatment information for breast cancer patients diagnosed in 1995. ~~Because of changes in patterns of health care delivery,~~ we are concerned that some treatment may be missing from our records. We realize that you may have already reported some treatment information to us or to your hospital cancer registry. However, where treatment information is missing from the consolidated abstract we submit to the CCR we are trying to determine whether it was not indicated and therefore not given, or if it was given and was not reported or included on the consolidated abstract.

We realize there are many demands of your time, so we have done all we can to keep the work required to a minimum. We are only requesting information for a small number of patients and treatments (i.e., those where there is a possibility, according to the National Cancer Institute Physician's Data Query (PDQ), that treatment may have been given; or where treatment was recommended by you or another physician but it is unknown if it was given).

For the specified treatment(s) noted on the attached form, could you please provide the information and return the form to us. **If you prefer not to complete the attached form, please notify us and we will make arrangements to come to your office to complete the work.** We know that you receive many requests for information and we want to make this request as easy as possible for you.

Thank you for your help. If you have questions, or would like us to complete the form for you, please call Helen Sanderson at (510) 429-2538.

Sincerely,

Dee W. West, Ph.D.
Executive Director
Northern California Cancer Center

Enclosure

P.S. We have enclosed a new brochure that the CCR just printed regarding the registry reporting system that you may find of interest.

PHYSICIAN'S QUESTIONNAIRE

Patient's Name:

Birthdate:

04/05/1912

Date of Initial Diagnosis:

07/24/1995

Cancer: Breast

Physician:

Dr. John Saranto

CSS No:

PLEASE PROVIDE THE *FIRST COURSE* OF CANCER-DIRECTED THERAPY REQUESTED BELOW. ALSO, WE WOULD APPRECIATE RECEIVING THE LATEST FOLLOW-UP INFORMATION THAT YOU HAVE ON THE PATIENT. THANK YOU.

TREATMENT INFORMATION REQUESTED

Did this patient have a lumpectomy? ___0 NO ___7 REFUSED ___9 UNKNOWN
___1 YES DATE: ___/___/___

Did this patient receive radiation therapy?
___0 NO ___7 REFUSED ___9 UNKNOWN ___1 YES DATE: ___/___/___

Did this patient receive chemotherapy? ___0 NO ___7 REFUSED ___9 UNKNOWN
___1 YES DATE: ___/___/___ Agent(s):

FOLLOW-UP INFORMATION

___ Alive Date of last contact _____
___ Dead Date of death _____
County/State of death _____

If you believe this information to be incomplete, are there other physicians we could contact who may have further information on this patient?

Dr. _____ Address _____

Dr. _____ Address _____

Physician Signature _____ Date _____

APPENDIX V

WRIGHT WE & ALLEN ME. CALIFORNIA CANCER REGISTRY
ENHANCEMENT FOR BREAST CANCER RESEARCH.

Poster Presentation at the Department of Defense
Breast Cancer Research Program Meeting

Washington, D.C.
November 1-4, 1997

California Cancer Registry Enhancement For Breast Cancer Research

**William E. Wright, Ph.D. & Mark E. Allen, M.S.
California Department of Health Services**

**Supported by the U.S. Army Medical Research and material Command
under DAMD 17-94-J-4508**

California Cancer Registry Enhancement for Breast Cancer Research

Abstract

The purpose of this project is to enhance the value of the California Cancer Registry (CCR) as a breast cancer research tool for clinicians and epidemiologists. The CCR began statewide population-based coverage on January 1, 1998. Between 1988 and 1993 all breast cancers were staged according to the National Cancer Institute's (NCI) Surveillance, Epidemiology, and End Results (SEER) Program Summary Staging Guide, basically a classification of cases into *in situ*, localized, regional, and distant disease. A major objective of this project has been to reclassify all breast cancer cases diagnosed between 1988-1993 according to the SEER Program's Extent of Disease classification scheme and to apply a computer program available from the NCI to classify cases into the TNM classifications and Staging Categories (0, I, II, III, IV) of the American Joint Committee on Cancer. This will allow for classification of breast cancer according to all staging schemes currently in use in the United States so that researchers could classify breast cancer cases according to the scheme most useful to their research.

A second objective of this project is to enhance the availability of breast cancer treatment data included in the CCR. Detailed and complete treatment data for all breast cancer cases is difficult to ascertain due to the fact that much treatment, especially chemotherapy, is given outside of acute care facilities. The approach has been to compare for individual patients the treatment information currently recorded into the data base to that recommended in the NCI's Patient Data Query (PDQ) data base. For all patients not recorded as having received the recommended treatment, follow-back to the physician of record occurs and he/she is queried regarding any additional treatment which the breast cancer patient may have received as a part of her initial course of therapy.

A third objective of this project has been to link the CCR breast cancer cases against other available data bases to enhance survival data by updating current vital status of breast cancer patients.

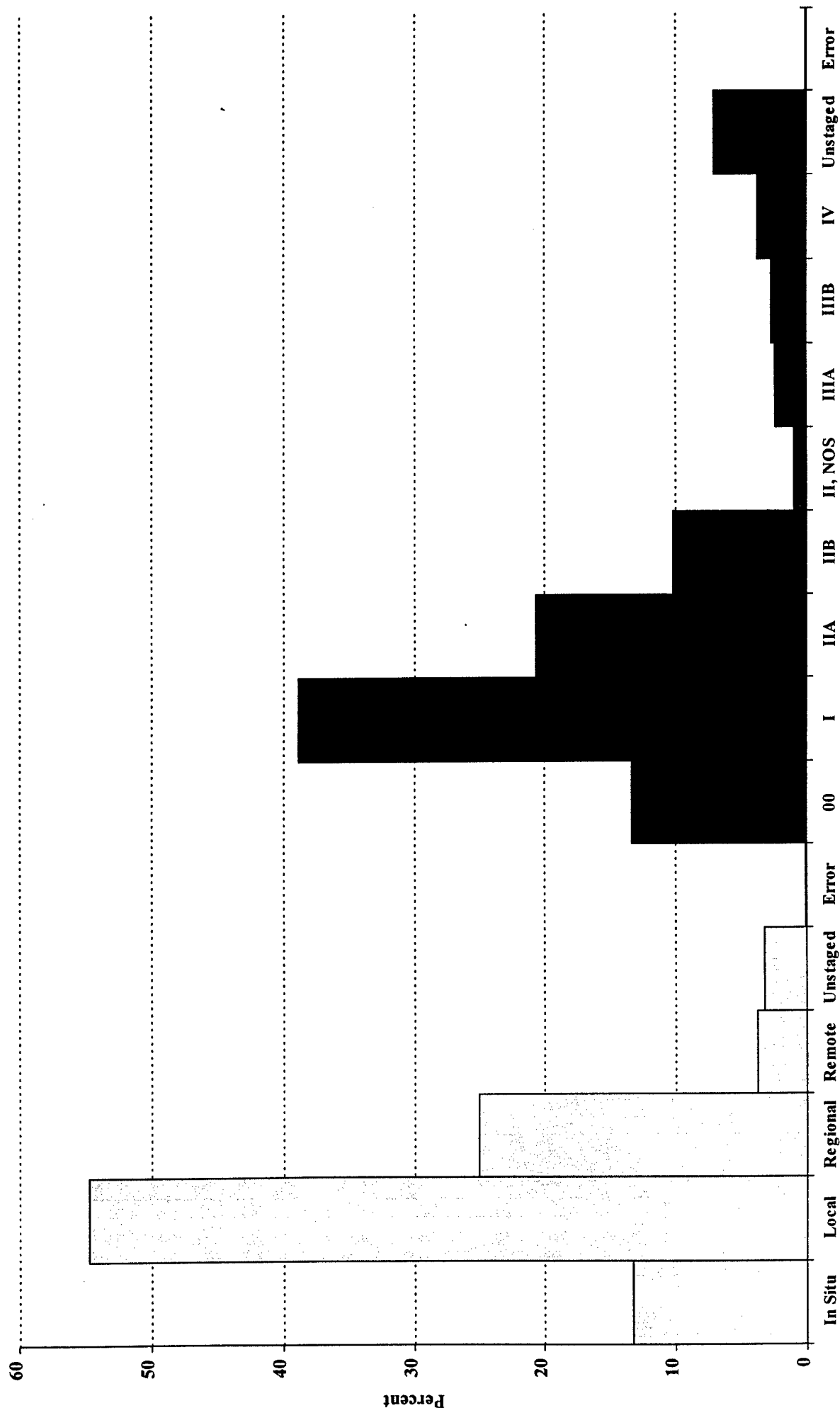
Currently, EOD coding has been completed for all of the 139,262 female breast cancers diagnosed in California between 1988-1994. (All cases diagnosed after 1/1/94 are required to have EOD codes reported to the CCR.) Software applications to convert these codes to AJCC Staging Categories have been completed. Physician follow-back for additional treatment data has been delayed due to a need for software to process a backlog of "correction" records. We determined that there is

an unknown amount of treatment information contained in the correction records that must be processed before any given breast cancer record could be compared with the standard recommended treatment and before follow-back to physicians concerning possible incomplete treatment could be initiated. Computer software to process correction records has been developed, and processing the backlog will begin in July, 1997 so that physician follow-back can commence this fall. A pilot test of physician office follow back has been completed for 1994 cases diagnosed in the San Francisco Bay Area Region of the CCR and additional treatment information was collected on 20% of the cases. Linkages with the Health Care Financing Administration (HCFA) MediCare files, hospital discharge files from the California Office of Statewide Health Planning and Development, and with the California Department of Motor Vehicles have been performed and vital status has been updated for over 20% of the breast cancer cases.

Extent of Disease Coding for Breast Cancer Cases, Diagnosed in 1994, Converted to SEER Summary Stage and AJCC TNM Staging

SEER SUMMARY STAGE

AJCC TNM STAGE



Breast cancer cases, diagnosed between 1988 and 1993, have been reclassified according to the SEER Program's Extent of Disease (EOD) classification scheme. Beginning in 1994, EOD staging will be collected for all cancer cases. EOD coding allows for the conversion to SEER summary stage, American Joint Committee on Cancer (AJCC) TNM stage and other classification schemes. This flexibility allows the researcher to choose the coding system which is most useful. For example, a researcher may choose SEER summary stage which designates a breast cancer into one of four broad staging categories or choose AJCC stages which has 8 categories, allowing for more detailed information to be recorded.

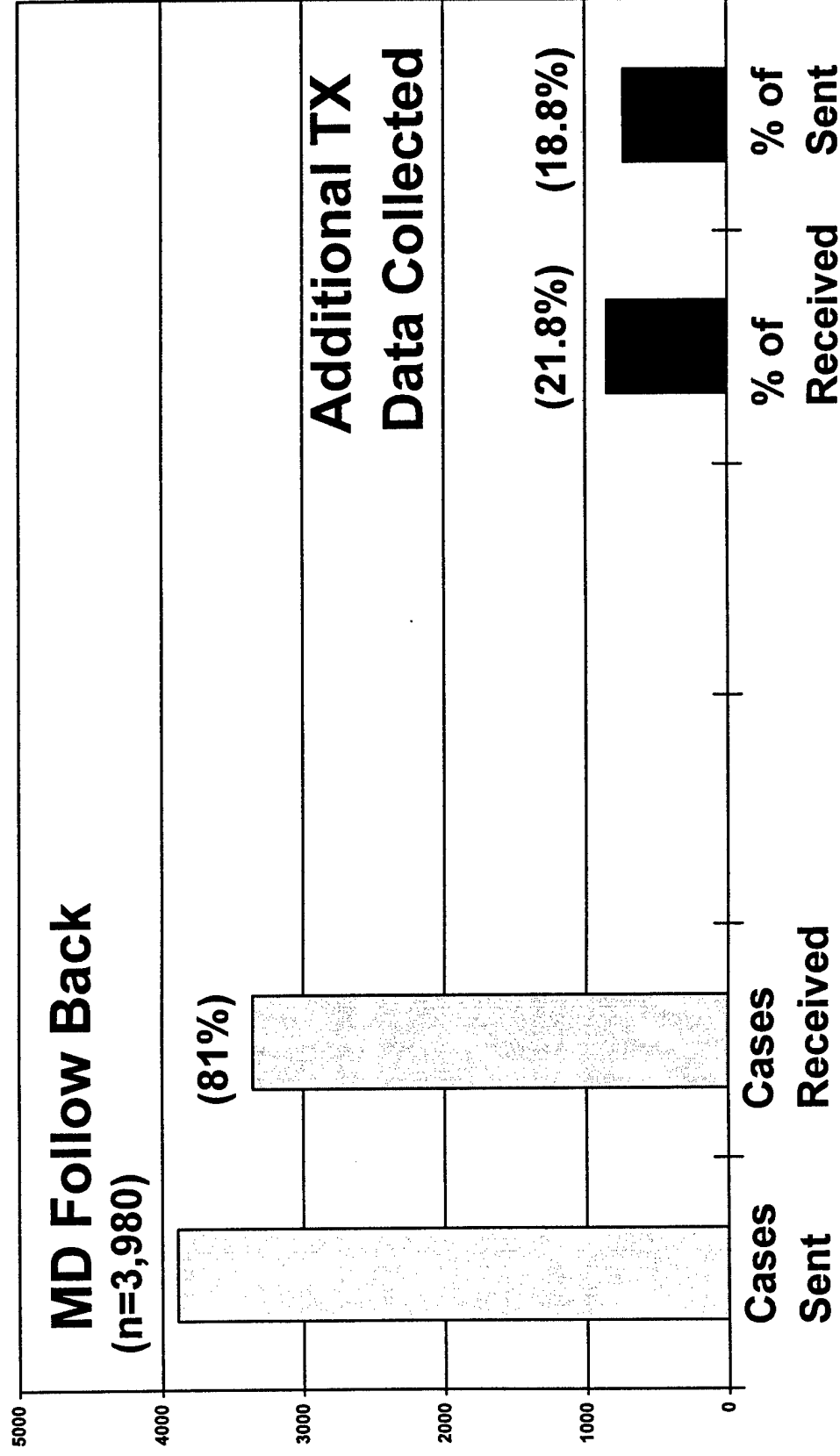
TREATMENT DATA AUGMENTATION

The San Francisco Bay Area Regional Registry of the CCR, which is also a component of the NCI's SEER program, conducted a study of the treatment information contained in their files for 4,795 breast cancer cases diagnosed in 1994. For each case, the combined first course of treatment (surgery, radiation, chemotherapy, hormone therapy) was compared with the NCI's Physician Data Query (PDQ) recommended breast cancer treatment guidelines for specific TNM stages using computer programs developed for this purpose. Only 336 (7%) of the breast cancer cases were recorded as having received the PDQ recommended first course of treatment.

Letters were written to physicians (or Kaiser hospitals for 1,053 cases) requesting additional treatment information for cases not meeting the PDQ standard. Follow back was not performed on 569 cases for various reasons (e.g. no contact MD available, hospital closed, cancer-directed surgery not recommended or contraindicated, etc.). The response rate was 86%, and additional treatment information was collected on 731 cases which is 21.8% of the cases for which responses were received and 18.8% of the total number that were queried. Radiation therapy which is usually administered in outpatient facilities was the most frequently reported additional treatment item.

Hospital and population-based cancer registries must implement follow-up procedures in order to have complete first course of treatment information on breast cancer cases.

1994 Treatment Data Augmentation Results of Physician Follow Back



Follow-up of Alive Women with Breast
Cancer, Diagnosed 1988-1993, Using
Hospital Discharge Files from 1991-1994

Current Follow-up Year on the Registry	Number of Women with follow-up during this year	Number updated from 1991 Hospital Discharge file	Additional number updated from 1992 Hospital Discharge file	Additional number updated from 1993 Hospital Discharge file	Additional number updated from 1994 Hospital Discharge file
1988	2850	254	209	155	149
1989	4850	468	364	303	256
1990	5752	653	420	350	332
1991	6944	482	700	481	420
1992	7107	0	487	654	552
Total	27497	1857	2180	1943	1709

114,010 women diagnosed with breast cancer in 1988-1993 were linked with the hospital discharge files for women aged 20 or over with a non-pregnancy related diagnosis in 1991-1994 from the California Office of Statewide Health Planning and Development (approximately 1.1 million records per year). The results from the above table indicate that 10% of the records may be updated for the first year of the linkage and 6-7% may be updated with subsequent years.

Comparison of Follow-up Results from Linkages with Different Data Bases

Data Base	Percent Updated from one year of data
Driver's License (Department of Motor Vehicles)	20
Health Care Financing Administration (HCFA) MediCare Files	0.3
Medicaid Files	8
Hospital Discharge Files	10

Linked 11,100 breast cancer cases with a diagnosis over one year old and a follow-up date over one year old with driver's license files and were able to update about 2,200 or 20%.

Linked 6,200 Health Care Financing Administration (HCFA) requests for payment in 1993 with 114,000 women diagnosed with breast cancer in 1988-1993. Only about 40 of the 27,500 (0.3%) alive women with a follow up date prior to 1993 could be updated.

Linked 1.15 million claims for payment by Medicaid in 1993 and were able to update 2,200 of 27,500 (8%) alive women with a follow up date prior to 1993.

PRESENTATIONS AND PUBLICATIONS

Allen ME, Wright WE, Perkins CI. Estimation of Case Ascertainment by Linkage with Medicare Files. Presentation at the Annual Meeting of the North American Association of Central Cancer Registries. Minneapolis, MN, April 17, 1996.

Allen ME, Wright WE, Perkins CI. Stage at Diagnosis and Treatment Among Medi-Cal Breast Cancer Patients in California. Presentation at the Annual Meeting of the North American Association of Central Cancer Registries. Boston, MA, April 2, 1997.

Parker JD, Newman J, Gegretsadik T, Kileen M. Managed care and treatment for early stage breast cancer: California Medicare, 1993. Medical Care (in press).

Morris CR. Comparison of SEER Staging and AJCC Staging for Breast Cancer.(Manuscript in preparation.)

CONCLUSIONS

- Population based cancer registries are an important tool for research on breast cancer causes and cures.
- Clinically relevant staging can be obtained by applying existing computer programs to information in cancer registry files.
- Cancer registry followup for survival analysis can be performed with computer linkages to Medicaid and Medicare files, hospital discharge files, and motor vehicle registration and driver's license files.
- Complete first course of treatment information for breast cancer is not currently available from hospital records. Physician followup must be conducted.
- Additional funding from the national and state level must be secured if population-based cancer registries are going to achieve their potential for serving research in breast cancer.

California Cancer Registry Enhancement for Breast Cancer Research

Summary

The purpose of this project is to enhance the value of the California Cancer Registry (CCR) as a breast cancer research tool. The CCR began statewide population-based coverage on January 1, 1998. Between 1988 and 1993 all breast cancers were classified into four categories based on the progression of the cancer: *in situ*, localized, regional, and distant disease. A major objective of this project has been to reclassify all breast cancer cases diagnosed between 1988-1993 into a different classification scheme developed by the American Joint Committee on Cancer. This will allow for classification of breast cancer according to all staging schemes currently in use in the United States so that researchers could classify breast cancer cases according to the scheme most useful to their research. A second objective of this project is to increase the amount treatment data on the breast cancer cases included in the CCR. A third objective of this project has been to link the CCR breast cancer cases with hospital and motor vehicle records bases to update vital status of breast cancer patients so that survival times can be computed.

Work on this project is proceeding. The project has generated two presentations to a professional society and one paper has been accepted for publication in a scientific journal. When completed, a data file with detailed information on breast cancer cases in California will be made available to qualified researchers. Future research using this file may lead to a better understanding of the causes and cures of breast cancer.